Topic: Relations

This paper is based on a first draft on Topic: Relations serving as discussion paper for the EAC-CPF meeting in Berlin, 9 - 12 March 2020.

1.   Introduction
Semantic Web / Linked Open Data

Today, we use links to identify entities (URI) and to locate information resources about entities on the internet (URL). This is an essential part of the semantic web, which builds on linked open data[1].

Linking entities and encoding relations between entities

When it comes to relations in the context of EAC-CPF and EAD, it is primarily suitable to distinguish between the aforementioned (semantic web) linking mechanism and the encoding of relationships. The last does not necessarily represent values as links but describes the nature of a relationship between (at least) two entities.

Therefore, the next part will summarize relationship types encoded with EAC-CPF and EAD and it discusses elements that constitute the description of relationships.

The subsequent part will focus on encoding relationships with EAC-CPF. It is the objective to work out both missing and redundant elements.

The last part is a draft recommendation for encoding relationships with EAC-CPF to promote its major revision. There are some simple objectives:

- Simplification of encoding while enhancing support for linked data
- Ensuring cross-domain interoperability
- Reconciliation efforts concerning both EAC-CPF and EAD

2.   Relationships in EAC-CPF and EAD
General Relationship Types

Relations are a crucial part in the description of archival resources[2] and the (social) context of agents who created and used the records[3]. In context of archival description we can distinguish:

- relationships as part of the archival description (agents, time, places, etc. identified with the description of record resources) and

---

[1] Tim Bernsers-Lee: Linked Data. https://www.w3.org/DesignIssues/LinkedData.html, 2006
[2] Wisser, Katherine et al.: Exploring Relationship Description. A Report from the Describing Relationships Workshop. Simmons College, February 2018. In: Journal of Contemporary Archival Studies. 6 (2019)
[3] Records in Context – A Conceptual Model. Consultation Draft v0.1, 2016

- relationships as part of the creation of metadata and authority records (agencies who have created, modified, … maintained the records)

Relationship types in the context of archival description

When digging into detail, EAS standards provide elements to encode relations to entities, e.g. //ead:persname|geogname|… or //eac-cpf:places|mandates|… which establish various relationship types:

EAC-CPF (relationship types)	EAD (relationship types)
CPF to CPF	Resource [4] to Resource
CPF to Function	Resource to Function
CPF to Resource	Resource to CPF (Origination, Controlaccess, Index, Custodhist, Relations)
CPF to Time	Resource to Time
CPF to Place	Resource to Place
CPF to Subject (Mandate, Function, Occupation, Legal Status)	Resource to Subject (Subject, Function, Occupation, Legal Status, Works)
CPF to Local Description	Resource to Works (Monographs/Periodicals) (bibliography)
	Resource to Instance (dao)

The table demonstrates similarities of EAC-CPF and EAD. It is thus encouraging to further also reconciliation efforts, which is, however, not the topic of this paper.

Before continuing, it is to the best advantage to discuss components constituting the description of relationships from a normative perspective first:

Components of a statement about an entity-relationship

Encoding a relation or any other property of an entity is primarily like encoding a statement: The mark-up for encoding structured data often provides a syntax to infer statements, e.g. "Fonds X was created by agent Y".

The quality of statements can differ significantly, from very general, e.g. "Person A is associated with Person B", to very contextual, e.g. "Person A has corresponded with Family B on date YYYY-MM-DD while at place X".

Proposal: Statements should be documented by a source, e.g. "Machine X has aggregated a statement on YYYY-MM-DDTHH:MM:SS+00:00 from source Z1. Person X modified it on YYYY-MM-DDTHH:MM:SS+00:00, based on source Z2" (s. paper on Topic: Assertions).

---

[4] Resource is defined as archival source being described; for published resources see works

Deductively, we can assume that statements about relationships are constituted by the following basic data components:

- Name of the related entity (mandatory)
- Type of the related entity
- Type of the relationship
- Role of the entity in a relationship
- Time or time frame of a relationship
- Place of a relationship
- Source of the relationship statement

Literals and Elements of Referencing a Linked Resource

To form at least a general statement about a relationship between two entities, we need to encode the name (e.g. part, term, *Entry) of the related entity. All components have in common that values can be expressed as literals.

It is, however, also desirable to encode the URI representing the literal (value) as well as both the authority / vocabulary source and its URI:

- Name of the authority / vocabulary source (literal)
- URI of the authority / vocabulary source (resource/URI)
- URI of the value (resource/URI)

These elements should be considered for encoding any reference for any value that can be encoded by both EAC-CPF and EAD.

Comparison

EAC-CPF provides a set of elements and attributes to encode relations:

EAC-CPF:description

As child of //eac-cpf/description:

- //function
- //legalStatus
- //mandate
- //occupation
- //place
- //localDescription

with child elements //citation, //date, //dateRange, //dateSet, //descriptiveNote, //placeEntry, and //term.

These elements can be used, with possibly only a //term child element in it, for indexing (categorizing using a controlled vocabulary), e.g. the function performed by a CPF entity.

Distinctions of //place

All elements share the same set of child elements and attributes, except //place with additional child elements:

- //place/placeEntry (vs. term)
- //place/placeRole
- //place/address

The child element //placeEntry has few more attributes assigned compared with //term:

- @accuracy
- @countryCode
- @localType
- @longitude

Both //placeEntry and //term encode the name of a related entity and share same attributes:

- @lastDateTimeVerified (date or time a linked resource was verified)
- @scriptCode (ISO 15924)
- @transliteration (naming the transliteration scheme)
- @vocabularySource (naming the controlled vocabulary source by its URI)
- @xml:id
- @xml:lang

The six parent elements (function, mandate, etc.) can be considered as subjects describing a CPF entity which usually do not provide specific information about role, place, or time.

While the aforementioned elements are child elements of the description part of EAC-CPF, the part //relations is dedicated for encoding relations between

- CPF-entities and CPF-entities (//cpfRelations)
- CPF-entities and functions (//functionRelations)
- CPF-entities and resources (//resourceRelations)

These three wrapper elements share the same set of child elements and attributes to specify relationships:

- relationEntry (name of the related entity)
- placeEntry (place of a relationship)
- date|dateRange|dateSet (time or time frame of a relationship)
- descriptiveNote (could be: Source of the relationship statement)
- objectXML|objectBinWrap *(other encoding schemes)*
- @[xlink:]actuate *(link behaviour)*
- @[xlink:]arcrole (Type of the relationship/nature of the relation)
- @[xlink:]href (URI representing the value)
- @[xlink:]role (URI of the type or nature of the related resource)
- @[xlink:]show *(link behaviour)*
- @[xlink]:title (name of the related entity / term of a linked vocabulary)
- @[xlink]:type *(type of linking, EAC-CPF: term "simple" only)*
- @[xml]:id *(document internal referencing)*
- @[xml]:lang *(language of the value)*

The relation elements use a third element to label the name of a related entity (relationEntry). The elements also provide a broader set of attributes to encode the nature of a link/reference compared

with //description/function|… etc. It is also providing additional elements for encoding the time and place of a relation.

The relations elements, when used, implies that a full description of the CPF entity, resource or performed function is hold or know elsewhere.

3. Recommendation

EAC-CPF is already providing a whole set of elements to describe various types of relations of CPF entities. Nevertheless, the encoding of relations and references is improvable for it is currently complex, sometimes ambiguous, and improvable.

The complexity is further increased by the singular/plural issue (#21, #61), e.g. it is ambiguous when to use @localType (//occupations or //occupation). It is the same with @vocabularySource (issue of the examples in the Tag Library which encode both literals and URIs for authority / vocabulary sources).

Proposal: **Linking and referencing attributes**, particularly attributes encoding the authority / vocabulary source, its base URI, and the value URI, **should be made available to all EAC-CPF elements encoding values that can represented by URIs** (example below).

Proposal: In addition to that, the recommendations concerning assertion description should be considered also for describing the relations, i.e. **enable the referencing attributes for the elements in the relation section**.

Finally, yet importantly, we should take the chance of a major revision to reduce the set of elements and attributes like actuate, type, id, … to focus on elements and attributes encoding, not formatting or (instructing) the processing of data.

It is recommended to draft best practice encoding principles for relations. The list of elements constituting relationships can be a starting point. Approaches of data provider like SNAC or GND should be examined for interoperability.

Proposal: The encoding of relationships should have the following elements available for a set of entity types (cpf|function|resource) within the wrapper element //relations:

- //relation (mandatory)
    - optional attributes @sourceref, and @maintenaceeventref
    - general optional attributes @audience, @encodinganalog, @id, @relatedencoding
- //relation/targetentity (mandatory) - *was //relation/relationEntry*
    - mandatory attribute @targettype (limited values: corporatebody | person | family | resource | function)
    - optional attributes @vocabularysource, @vocabularysourceuri, @valueuri
    - general optional attributes @audience, @encodinganalog, @id, @relatedencoding
- //relation/targetentity/part (mandatory)
    - optional attribute @localtype
    - general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration
- //relation/relationtype - *new element*
    - optional attributes @localtype, @vocabularysource, @vocabularysourceuri, @valueuri

- o general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration
- · //relation/targetrole - *new element*
    - o optional attributes @localtype, @vocabularysource, @vocabularysourceuri, @valueuri
    - o general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration
- · //relation/date|daterange|dateset
- · //relation/place
- · //relation/place/targetentity
    - o mandatory attribute @targettype (limited values: corporatebody | person | family | resource | function)
    - o optional attributes @vocabularysource, @vocabularysourceuri, @valueuri
    - o general optional attributes @audience, @encodinganalog, @id, @relatedencoding
- · //relation/place/targetentity/part
    - o optional attribute @localtype
    - o general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration
- · //relation/place/relationtype
    - o optional attributes @localtype, @vocabularysource, @vocabularysourceuri, @valueuri
    - o general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration
- · //relation/place/targetrole
    - o optional attributes @localType, @vocabularysource, @vocabularysourceuri, @valueuri
    - o general optional attributes @audience, @encodinganalog, @id, @lang, @relatedencoding, @script, @transliteration

Example (Paul Arendt, parent of Hannah Arendt)

```
<cpfRelation cpfRelationType="family" xlink:arcrole="https://d-nb.info/standards/elementset/gnd#Family"
xlink:href="https://d-nb.info/gnd/118561715">
    <relationEntry>Arendt, Paul (1873–1913) [Eltern]</relationEntry>
</cpfRelation>
```

As it would be better:
```
<relations>
    <relation id="1" sourceLink=#43" maintenanceLink="#43">
      <targetEntity vocabularySource="GND"
      vocabularySourceURI="http://d-nb.info/gnd/"
      valueURI="http://d-nb.info/gnd/1190671301" targetType="Person">
        <part localType="family">Arendt</part>
        <part localType="given">Paul</part>
        <part localType="dates">1873-1913</part>
      </targetEntity>
      <relationType voabularySource="GND"
      vocabularySourceURI="https://d-nb.info/standards/elementset/gnd#"
      valueURI="https://d-nb.info/standards/elementset/gnd#Family" localType="text"
      lang="de">Familie</relationType>
      <relationType voabularySource="GND"
      vocabularySourceURI="https://d-nb.info/standards/elementset/gnd#"
      valueURI="https://d-nb.info/standards/elementset/gnd#Family" localType="text"
      lang="en">Family</relationType>
      <relationType voabularySource="GND"
      vocabularySourceURI="https://d-nb.info/standards/elementset/gnd#"
```

```
            valueURI="https://d-nb.info/standards/elementset/gnd#Family"
            type="code">bezf</relationType>
         <targetRole voabularySource="schema.org"
         vocabularySourceURI="http://schema.org" valueURI="http://schema.org/parent"
         localType="text" lang="de">Eltern</targetRole>
         <targetRole voabularySource="schema.org"
         vocabularySourceURI="http://schema.org" valueURI="http://schema.org/parent"
         localType="text" lang="en">parent</targetRole>
         <targetRole voabularySource="schema.org"
         vocabularySourceURI="http://schema.org" valueURI="http://schema.org/parent"
         localType="code">par</targetRole>
         <date>...</...>
         <place>...</...>
      </relation>
</relations>
```

Compared with MARC21 (GND):

```
<datafield tag="500" ind1="1" ind2=" ">
   <subfield code="0">(DE-588)1190671301</subfield>
   <subfield code="0">https://d-nb.info/gnd/1190671301</subfield>
   <subfield code="a">Arendt, Paul</subfield>
   <subfield code="d">1873-1913</subfield>
   <subfield code="4">bezf</subfield>
   <subfield code="4">https://d-nb.info/standards/elementset/gnd#familialRelationship</subfield>
   <subfield code="i">Beziehung familiaer</subfield>
   <subfield code="9">v:Vater</subfield>
</datafield>
```

It is to be discussed if the following elements and attributes should be available for //mandates|occupations|etc., too.

4.   Summary

We should have only one element for //relations with only one, but repeatable, child element //relations/relation.

The element //relation should have an attribute @targettype to specify the type of the related entity (CPF, function, resource, …?). It should also have an attribute @id to encode any document-internal ID for internal referencing.

The element relation should have the following child elements:

- //targetentity (for labelling the name of the related entity)
- //relationtype (for labelling the type of the relation)
- //targetrole (for labelling the role of the related entity)
- //place (for labelling the place a relation is identified with)
- //date|… (for labelling the time a relation is identified with)

The element //targetentity should have a child element //part with an attribute @localtype (for specifying the value: family and given name, living dates, …).

The child elements //targetentity, //relationtype, //targetrole, and place (here: placeentry, placerole, …) should have the attributes:

- @vocabularysource (for labelling the authority or vocabulary source)
- @vocabularysourceuri (for encoding its base URI)
- @localtype (for encoding the value URI)
- @type (for providing information about the element value – text or code)

- @lang (for providing a language code for the element value)

The element //targetentity with //part should be mandatory.

A reconciliation of labelling the elements for encoding a name (term, placeEntry, or relationEntry) should be reconsidered.